Supporting Information for "Detecting mesopelagic organisms using biogeochemical-Argo floats"

Nils Haëntjens[1], Alice Della Penna[2,3], Nathan Briggs[4], Lee Karp-Boss[1], Peter Gaube[2], Hervé Claustre[5], and Emmanuel Boss[1]

[1]School of Marine Sciences, University of Maine, Orono, ME, USA.

[2]Applied Physics Laboratory, University of Washington, Seattle, WA, USA.

[3]Laboratoire des Sciences de l'Environnement Marin (LEMAR), UMR 6539 CNRS-Ifremer-IRD-UBO-Institut Universitaire Européen de la Mer (IUEM), Plouzané, France

[4]National Oceanography Centre, Southampton, United Kingdom.

[5]CNRS et Sorbonne Université, Laboratoire d'Océanographie de Villefranche, UMR 7093, 06230 Villefranche-sur-mer, France

Spike Layer Detection Algorithm

Algorithm Description

The algorithm to detect spike layers consists of two major steps: first finding individual spikes in profiles and secondly identifying if the spikes form layers. The algorithm is applied only on sections of the profile with a local sampling depth resolution greater than 1 observation/5 dBar (even if no spike layers were found in profiles with a sampling resolution lower than $\sim$ 1 observation/2 dBar, Figure S1).

A 15 point Hampel filter (similar to median filter except it only affects outliers) is applied recursively to x, the profile of the optical property of interest (e.g., FDOM, b_{bp}, or fchl), to smooth the profile ($x_{filtered}$). Spikes (x_{spikes}) are considered if they meet the following criteria:

$$x - x_{filtered} > x_{err} \tag{1}$$

with x_{err} a threshold defining the minimum intensity of the spikes, this threshold is specific to the property of interest (provided in S1).

Spikes at depth >10 dBar are clustered together using hierarchical clustering with a euclidean distance cut off of 50 dBar (function cluster from Matlab MathWorks, 2019). For each cluster of spikes a set of features are computed and then used for the classification of the profiles: number of spikes (n), median pressure (p), thickness (t), spike density ($d = n/t$), and median normalized spike intensity ($i = median((x_{spike} - min(x))/median(x - min(x)))$. The clusters of spikes meeting the following criteria are considered a layer of spikes potentially caused by mesopelagic organism:

$$\begin{cases} n \geq N \\ T_{min} \leq t \leq T_{max} \\ i \geq I \end{cases} \tag{2}$$

With N the minimum number of spike per cluster (set to 3 for this study), T_{min} and T_{max} the minimum and maximum thickness of the spike cluster, and I the normalized intensity threshold of the spike cluster. The value of these parameters are given in Table S1. A Matlab implementation of this spike layer detection algorithm is available at `https://github.com/OceanOptics/FloatSpikeAnalysis/`.

Algorithm Performance

The performances of the spike layer detection algorithm was assessed for both FDOM, b_{bp}, and fchl channels with commonly used metrics: accuracy, precision (P), recall (R), and F1-Score (Sokolova & Lapalme, 2009). The validation dataset consists of 6889 profiles manually annotated: presence of spike layer (n=127) and absence of spike layer (n=5648), which are the only one used here. Two additional categories, which consist of dubious (n=72) and unusable (n=1042) were ignored. Dubious consists of profiles that could not be classified as present or absent, and unusable consists of shallow profiles (<100 m), under-sampled profiles (<50 observations), or unrealistic profiles (e.g. many negative spikes likely due to dysfunctional sensors).

The assessment of the algorithm (Table S1) suggest that the algorithm performs well for FDOM profiles. For b_{bp} profiles the precision could be high (>90 %) however many spike layer are missed by the algorithm. For fchl profiles the algorithm is not reliable for detection of mesopelagic organisms which is likely due to the fact that spikes in fchl profiles did not form distinguishable layers and that associated intensities are very small in many cases. The superior performance of the FDOM algorithm resides in the fact that b_{bp} profiles contains additional spikes due to aggregates making it harder to distinguish zooplankton associated spike layers, while FDOM seems not to be sensitive to those aggregates. A combination of the information contained in both b_{bp} and fchl did not improve the performance of the algorithm with respect to running it for b_{bp} alone. While the statistics of spike layer detection algorithm support using it with no further validation, especially with FDOM profiles, we nonetheless, recommend for ecological studies to validate manually validate all profiles identified as containing spike layers.

Float Sampling Resolution

To evaluate the sampling resolution required by the floats to detect spike layers, we first looked at the average distance between observations in each profile of the BGC-Argo database. Only the upper 1000 dBar of the profiles were considered. We then investigated the relative occurrence of profiles with spike layers concerning their sampling frequency. This analysis, presented in Figure S1, helped to make an informed recommendation on the minimum sampling resolution to detect spike layers in float profiles: >1 observation/2 dBar. The full spatial coverage of the BGC-Argo floats with backscattering sensors is shown in the map Figure S1.a, while, the coverage with adequate sampling resolution is shown in the manuscript (Figure 3.a).

NAAMES Acoustics

Three matchups of float profiles and ship-based acoustics of NAAMES (Table S2) exhibited spike layers in the FDOM and b_{bp} float profiles. The case of September 5-7, 2017 presented in Figure 1 and 2 is further illustrated with a timeseries of the acoustic at 120 kHz (Figure S2) which is better suited for observing relatively smaller organisms (1 cm) but limited to the upper 200 m (due to the increased attenuation of sound in that frequency). The timeseries of the acoustic at 120 kHz shows shows similar patterns as the acoustics at 38 kHz for the same period. Besides these matchups of September 5-7 2017 (Campaign 3, Station 2), similar acoustic patterns paired with float spike layers were observed on May 28-31, 2016 (Campaign 2, Station 5) and September 9-11, 2017 (Campaign 3, Station 4) are presented in Figure S3.

Experiment with zooplankton and FDOM sensors

To test if zooplankton can produce spikes in FDOM signals similar to the in situ observations and to test for phototactic behavior an experiment was set up in 3 phases with sensors similar to the one mounted on the floats (an ECO-FLBBCD and an ECO-FLCDRT).

The experiment was conducted aboard a research vessel at station Papa (north east Pacific), in September 2018, with zooplankton collected $\sim$ 2 hours earlier with a MOCNESS net during daytime. A black tub with a black cover was filled with $\sim$ 80 L of surface seawater and a first set of measurements was taken with both sensors to measure the background signal. Next, $\sim$2 L of highly-concentrated zooplankton was added to the tub and a second set of measurements was taken. Finally, the visible LED of the FLBBCD sensor were covered with black tape such that no visible light was emitted (to avoid inducing a behavioral response towards or away from the sensor) and a third set of measurements was taken (only the FDOM channel is expected to receive a signal). Attention was paid to minimize the tub effect on sensors reading (mainly plastics fluorescence). The tub was kept dark during measurements so that the only possible source of light inside the tub was that of the sensors.

During the first set of measurements (with no zooplankton) no spike were observed with both sensors. The baseline of FDOM slightly increased during the second set of measurements (when zooplankton was present) and spikes similar to the one observed on float profiles were observed.

If the zooplankton are attracted to visible light then we would expect the third set of measurements to have a significantly lower frequency of spikes than the second set of measurements. The frequency of spikes during the third set of measurements was similar to the second one, suggesting that the organisms (mostly *Neocalanus cristatus*, *Vibilia*, *Themisto pacifica*, and *Clausocalanus lividus*) used during this experiment did not manifest an attraction behavior to the ECO sensors. The lack of response could have been induced by the fact that the zooplankton were recently caught by nets or the species were different than those encountered by the floats with different phototactic sensitivity and behavior.

References

MathWorks. (2019). *MATLAB Statistics and Machine Learning Toolbox.*

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, *88*(424), 1273–1283.

Sokolova, M., & Lapalme, G. (2009, 7). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. doi: 10.1016/j.ipm.2009.03.002

Table S1. Spike layer detection algorithm parameters and associated performance to classify float profiles as containing (present) or not (absence) layers of spikes.

						Presence			Absence		
Channel	x_{err}	I	T_{min}	T_{max}	Accuracy	P	R	F1	P	R	F1
FDOM	3×sMAD*	1.2	3	350	94	68	95	79	99	94	97
FDOM	0.9	1.2	3	350	96	84	86	85	98	98	98
b_{bp}	3×sMAD*	10	10	350	85	40	63	49	95	88	91
b_{bp}	0.00479	10	10	350	94	91	52	66	94	99	97
b_{bp}	0.00281	10	10	350	94	80	65	72	95	98	97
fchl	3×sMAD*	1.2	3	350	82	35	65	46	95	84	89
fchl	0.04	1.2	3	350	83	37	67	48	95	85	90

*sMAD is the scaled median absolute deviation of x over the whole profile as defined by Rousseeuw and Croux (1993)

Table S2. Synthesis of spikes observed with NAAMES BGC-Argo floats collocated with acoustics observations from the R/V Atlantis within 8 days. The acoustic backscattering strength S_v is classified as: *(absent)* no evidence of DVM could be visually identified on a 24 hours period, *(weak)* DVM patterns are observed and surface S_v was smaller than -80 dB at night, and *(strong)* DVM patterns are observed and surface S_v was greater than -80 dB at night. The number of spikes per profiles was quantified as follow: *(single)* multiple distinct spikes along the profile, *(collocated)* multiple sets of up to three spikes collocated on the profile, and *(layer)* a layer of spike, more than 5 collocated spikes. Note that the three matchup stations with spike layers are presented in details in Figure 1 and S3. Δz corresponds to the mean distance between the depth of the FDOM spike and the DVM scattering layer, a negative sign means that the spike is deeper than the acoustic layer. All matchups profiles are ascending.

Campaign	Station	WMO	DVM	FDOM Spikes	b_{bp} Spikes	Δz (m)
1	2	5902462	Absent	None	None	—
1	4	5902460	weak	None	None	—
2	1	5903102	weak	None$^{\alpha}$	Single$^{\alpha}$	—
2	2	5902462	weak	Single	Single	200
2	4	6901180*	Strong	—$^{\beta}$	Single$^{\gamma}$	—
2	4	5903101	Strong	None	Single	—
2	5	6901525	Strong	**Layer**	**Layer**	0 ± 0
2	5	5903100	Strong	None$^{\alpha}$	Single$^{\alpha}$	—
3	1	5903109	Strong	None	None	20
3	2	5903108	Strong	**Layer**	**Layer**	30 ± 35
3	3	5903107	Strong	Single	Single	3 ± 5
3	4	5903106	Strong	**Layer**	**Layer**	50 ± 30
3	5	5903105	Strong	Single	Single	-20 ± 23
3	6	5903104	Strong	Collocated	Single	-6 ± 13
3	6	5903103	Strong	Collocated	None	-11 ± 21
4	1	5903109	Strong	None	Single	—
4	2	5903108	Strong	None	None	—
4	2RF	5903108	Strong	None	None	—
4	3	5903107	weak	None	None	—
4	4	5903106	weak	None	None	—

$^{\alpha}$ The float only profiles at dawn.
$^{\beta}$ The float has a channel of $b_{bp}(532)$ in place of FDOM.
$^{\gamma}$ b_{bp} spikes are spread along the entire profile.

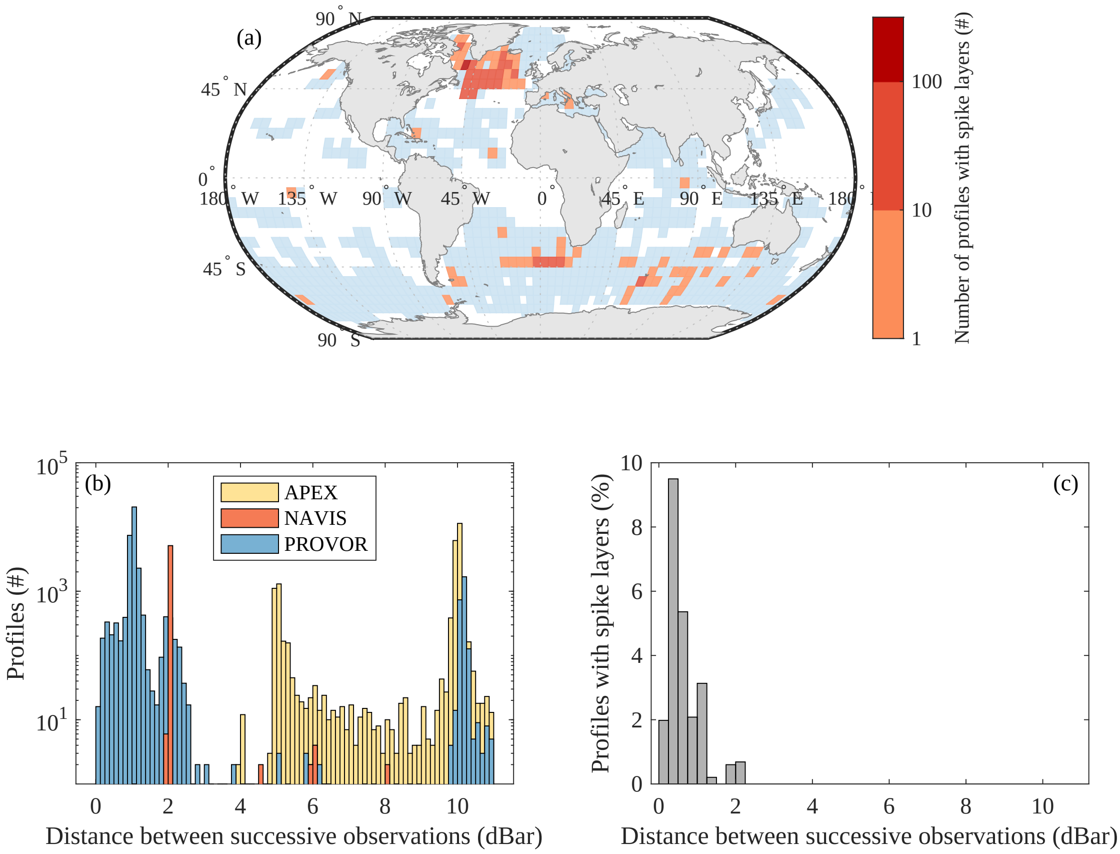

Figure S1. BGC-Argo Floats Sampling Resolution. (a) Locations of all BGC-Argo float profiles (no restriction on profiling resolution) in a 5×5° grid, blue areas indicate the presence of float profiles and red areas indicate the number of profiles with b_{bp} spike layers. (b) Number of BGC-Argo float profiles as a function of the median distance between successive observations within each profile, colored by platform type: APEX (yellow), NAVIS (red), and PROVOR (blue). Note that 3 % of the profiles with a median observing distance greater than 11 dBar are not shown. (c) Histogram of the number of profiles with spike layers normalized by the total number of profiles as a function of the median distance between successive observations within each profile.

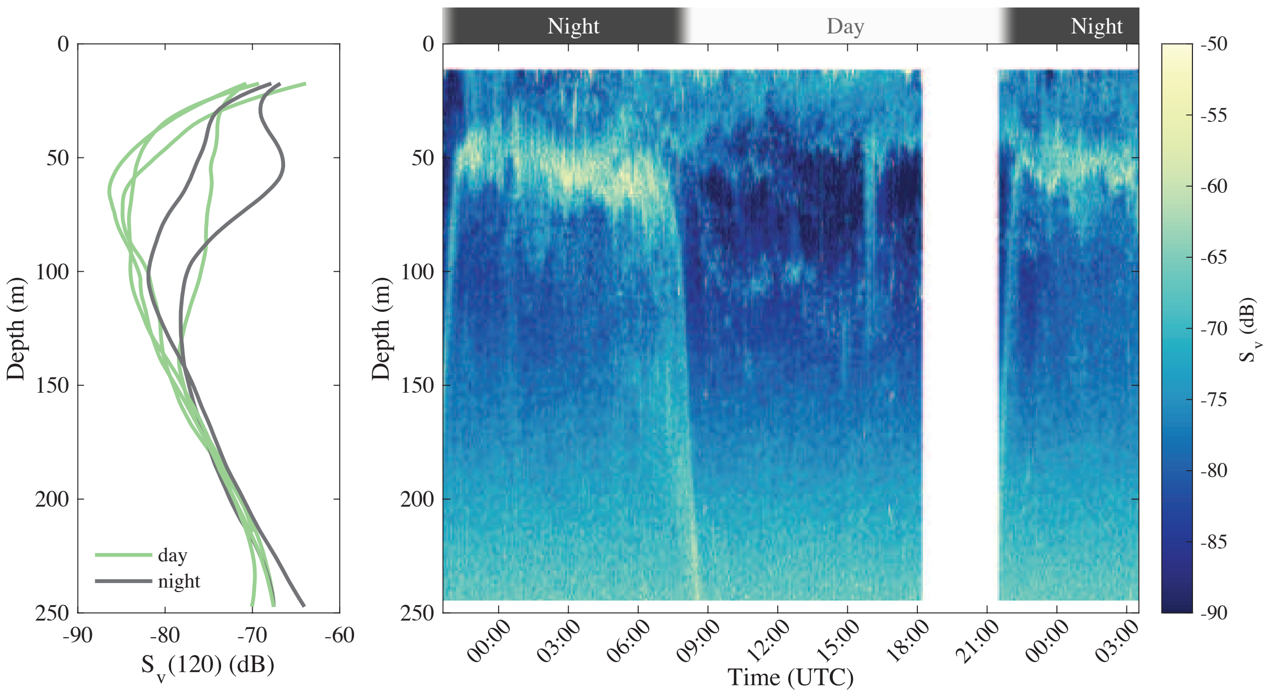

Figure S2. Profiles *(left)* and timeseries *(right)* of the mean volume backscattering strength at 120 kHz ($S_v(120)$) from the pole-mounted echosounder of the ship at the same time and location of $S_v(38)$ presented in Figure 1. Dark lines correspond to night-time profiles and green lines correspond to day-time profiles. The time of the acoustic profiles matches exactly with the time of the float profiles presented in Figure 2. Note that the higher profile of $S_v(120)$ during the day time was from 15:00 to 15:30, time during which $S_v(120)$ is temporarily higher. The first 10 m of acoustic data are cropped due to near-field effects. No acoustic backscatter was collected between 18:12 and 21:29.

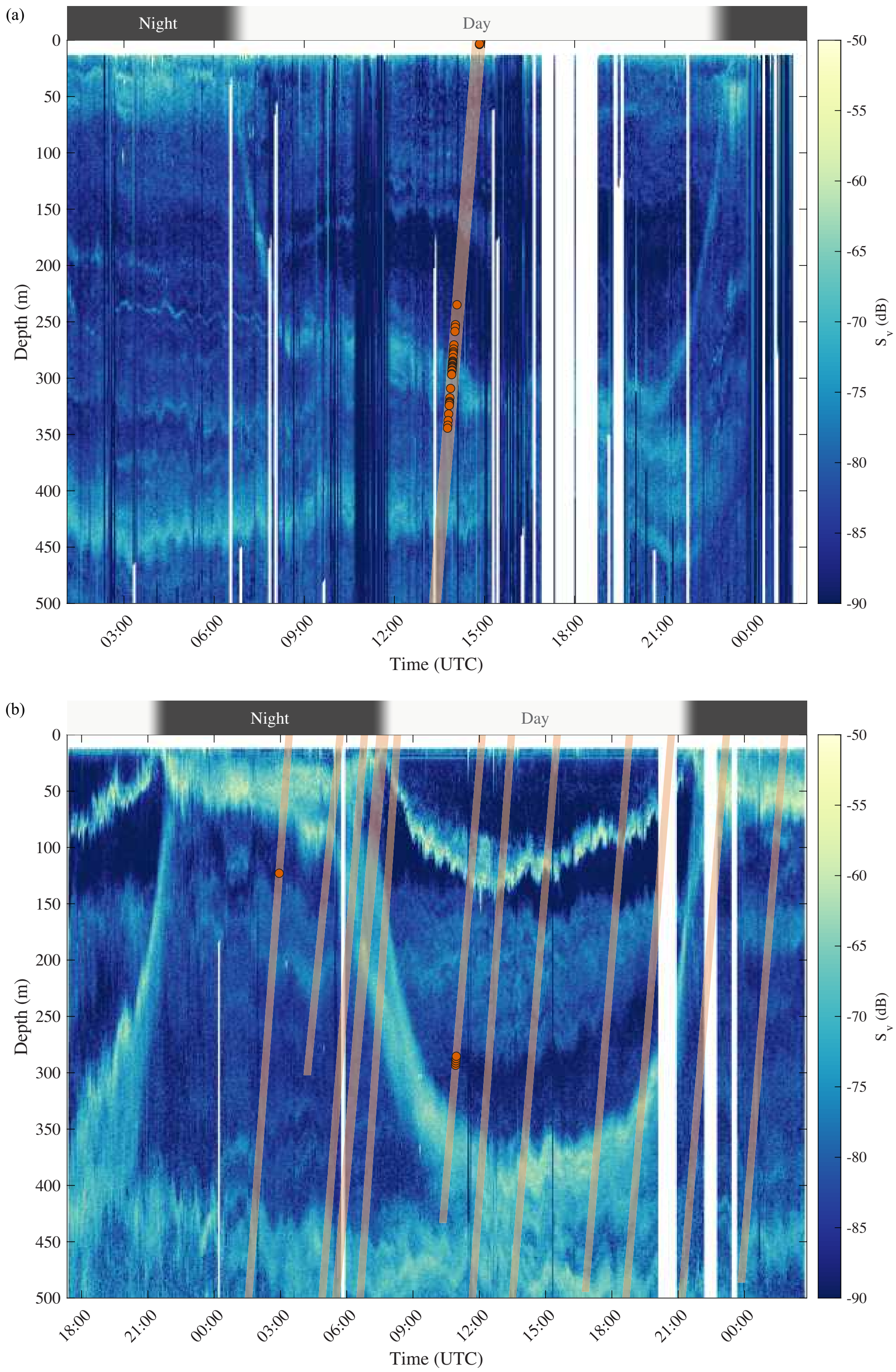

Figure S3. Time-series of the mean volume backscattering strength (S_v) at 38kHz at the second NAAMES campaign station three *(a)* and the third NAAMES campaign station four *(b)*. The orange lines are the up-casts of the float and orange circles superimposed on the profiles are FDOM spikes. The float profiles are within eight days of the acoustic. The slopes of the lines correspond to the profiling speed (0.08 m/s for the float). The first 10 m of the data are removed to mask near-field effects in the acoustics signal.